

Abstract

Modern neuroimaging methods have become a primary scientific tool for studying the human mind. One important question is how to deal with the relatively limited spatial and temporal resolution of neuroimaging data. Advanced statistical algorithms and computational modelling have helped to push the boundaries of this technique, but the scope of questions that can be addressed remains limited. Here, we seek to advance a novel analysis method, model-based multivariate fMRI analysis, to investigate one of the currently most pressing questions about human reinforcement learning: how is reward-irrelevant environmental information filtered from neural representations, to allow for efficient learning? The collaboration will involve the Niv lab at Princeton University, who are leading experts in model-based neuroimaging of reinforcement learning tasks, and the Haynes Lab at Humboldt Universität, who are at the forefront of development of multivariate fMRI analysis methods. Both PIs know each other, but have not collaborated before. Through the proposed collaboration, this project promises to lay the foundation for important theoretical as well as methodological advances, and for a long-lasting collaboration.

Summary

A central challenge in neuroscience is to understand the neural and computational basis of reinforcement learning (RL) – how animals and humans learn to make good decisions based on trial-and-error interaction with the world. The existing computational framework of RL has been extremely successful in explaining learning and decision making in the brain in *simple* environments (e.g., Niv and Schoenbaum, 2008; Niv, 2009). However a fundamental shortcoming of RL algorithms comes from the famous “curse of dimensionality” whereby learning slows down exponentially as the environment becomes more complex. How is it, then, that humans and animals can learn in complex environments with relative ease? What is missing from RL methods? The Niv lab has recently proposed a potential solution: throughout the learning process, neural representations are constructed that *filter out* environmental dimensions that are not important for the task at hand. This serves to reduce the complexity of the task and make it solvable with RL. We have developed computational models of how these filters are learned and shaped by experience, and have applied them to data from human participants learning to solve a multidimensional choice task (Wilson and Niv, 2012; Niv et al., in press). Our models make predictions about distributed representations in the cortex that differ substantially from those suggested by previous work. These predictions are exciting and novel, but at the same time challenging to test because current methods of analysis of human neuroimaging data usually focus on detecting scalar quantities, rather than distributed (so-called multivariate) representations.

Two methodological approaches are particularly relevant for this challenge: ***Model-based fMRI analysis***, on the one hand, is a method that can be used to test precise hypotheses about the temporal dynamics of fMRI signals that are derived

from computational models. This method has been frequently employed in the Niv lab (e.g., Niv et al., 2012, in press). Unfortunately, this method can only handle scalar (univariate) predictions (see Figure 1, upper panel). **Multi-voxel pattern analysis** (MVPA), on the other hand, to whose development the Haynes lab has made important contributions (e.g., Haynes and Rees, 2005; Kahnt et al., 2010), addresses questions about the informational content encoded in distributed fMRI signals (See Figure 1, middle panel). Currently this is only possible for stationary signals of interest, not for the dynamically changing representations that are at the core of our models of RL in complex, real-world decision-making. The Haynes lab, however, has recently started developing a more general method for analysis of fMRI signal patterns that addresses these problems (Allefeld and Haynes, 2014).

The goal of this project is to initiate a collaboration between the Haynes and Niv labs, that will allow us to develop specific analysis methods for fMRI data, which can track distributed representations that change over time. Towards this goal, we have two specific aims in the immediate timeframe:

- 1) ***To apply the Haynes' lab recently-developed cvMANOVA multivariate fMRI analysis approach to our existing data, in order to identify task-relevant distributed representations in the brain.***
- 2) ***To develop a novel, model-based multivariate fMRI analysis method that extends cvMANOVA to time-varying representations and can be used as a better tool to test our computational model of RL.*** This new method will be tested on the same data as in Aim 1, in order to compare and contrast the strengths and weaknesses of the two methods.

We hope to achieve these aims by synthesizing the expertise in the lead applicants' labs in computational models of RL and in fMRI analysis methods. The results of this project will not only inform the development of the next generation of computational RL models, but they will also provide analysis tools broadly usable by the large community studying learning and decision making in the human brain. ***The project therefore promises to answer important questions about how humans learn from rewards and to expand the scope of questions that can be tested with fMRI more generally.***

Background

Studying reinforcement learning using model-based fMRI analysis

Building computational models of neural processes to understand behavior has been a vastly successful approach. The poster child of this endeavor are models of RL, which have proposed that the difference between predicted reward and actual reward—the so-called “reward prediction error”—is an important quantity for learning from trial and error (Sutton and Barto, 1998). Strikingly, subsequent research has identified signals corresponding to prediction errors and learned values in the brain, and has shown that these signals relate to overt learning and choice behavior (Niv, 2009; Glimcher, 2011). In particular, fMRI studies

investigating the substrates of RL and decision-making have used a “model-based” fMRI approach: the dynamics of learning in the task are first expressed as a computational RL model, which embodies the hypothesis to be tested. Precise predictions about the development of prediction errors and values throughout the task are then extracted from this model and tested against fMRI data (see Figure 1, upper panel). This allows researchers to search for neural signals that correspond to “internal” variables postulated by a model, and to gain insight into how different brain areas coordinate the computations underlying behavior.

Following initial successes of this methodology, however, recent human RL research has increasingly emphasized some major outstanding questions regarding this learning process (e.g., Dayan and Niv, 2008). In particular, it has been challenging to move away from simplified learning situations in the lab, and cash in on the promise of RL theories to explain real-world learning and decision making in complex, ever-changing environments. This is because computationally, RL methods scale very badly to complex, multidimensional tasks – training time increases exponentially with the number of dimensions of the environment, making learning in even moderately complex environments hopelessly inefficient. To explain how it is that humans and animals can nevertheless learn new tasks as rapidly as they do, we have proposed that RL mechanisms in the brain operate on highly filtered neural signals, which effectively reduce the dimensionality of the problem to a tractable level (Gershman and Niv, 2010; Niv et al., in press; Wilson and Niv, 2012). This reduction entails representing only the aspects of the environment that are relevant for obtaining reward. One intuitive implementation of such dimensionality reduction is through mechanisms of *selective attention*: aspects of the environment that are inconsequential to task performance and attainment of reward can be ignored, while those that are relevant for the task at hand should be attended to. ***A central question is, therefore, how does the brain determine what information to attend to during reinforcement learning?***

We have already begun to develop extensions of classical RL models, which specify how attentional filters change during learning in multidimensional environments (Niv et al., in press; Wilson & Niv, 2012). While our research shows that accounting for the changes in attention leads to better predictions of behavior and mean neural activity (Niv et al., in press), it is not yet clear whether the assumed changes in neural representations can be identified in and studied using fMRI signals. What makes this endeavor challenging is that such representations are distributed, multivariate signals, while existing model-based fMRI analyses are only applicable for testing predictions about univariate signal changes. ***The lack of methods for testing multivariate model-based predictions in fMRI signals is thus a major impediment to further understanding how RL can operate in complex environments. The main goal of this project is to overcome this methodological hurdle and develop multivariate model-based fMRI methods.***

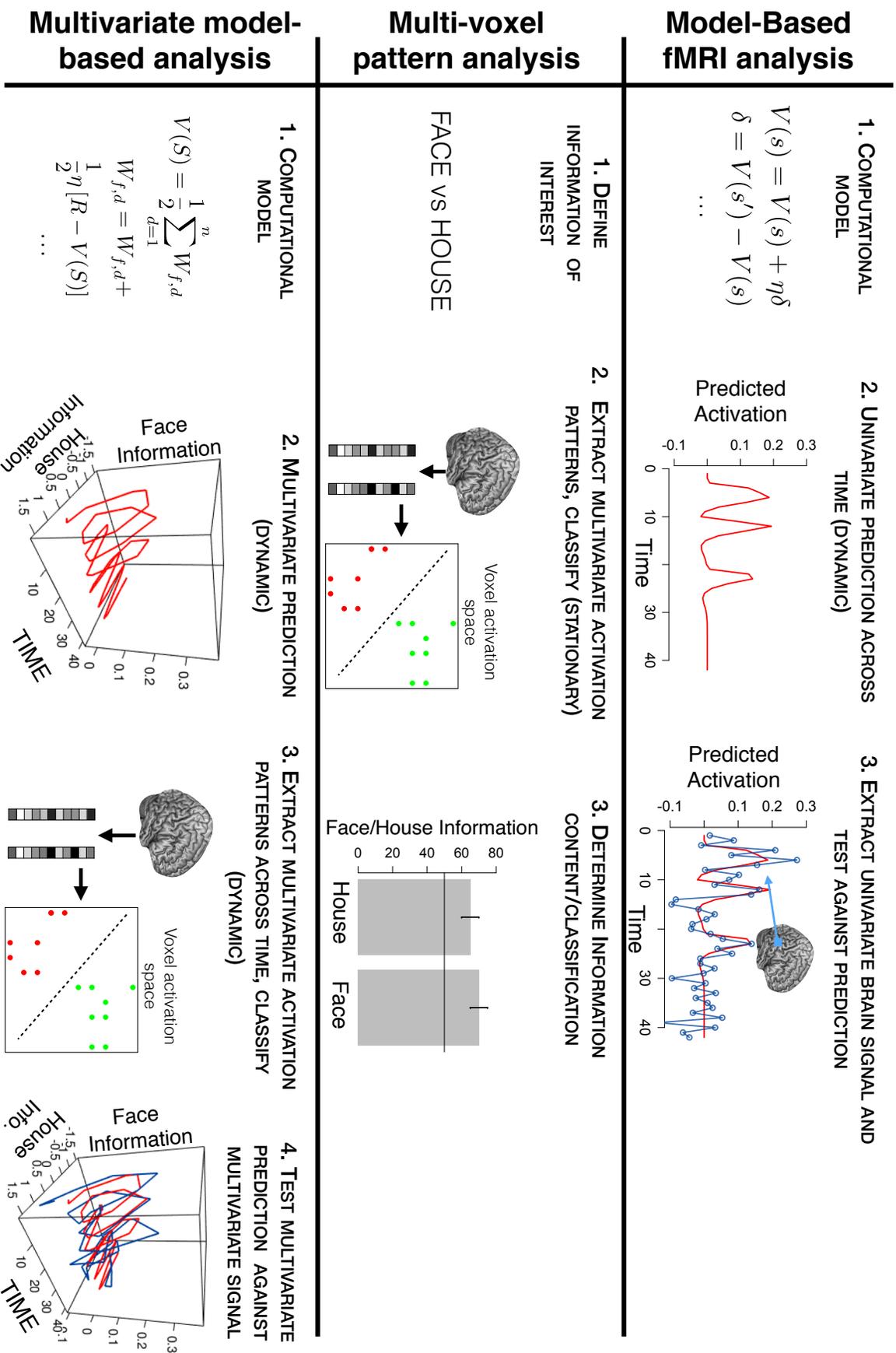


Figure 1. Summary of existing and proposed analysis approaches. Upper panel: Model-based fMRI analysis uses univariate brain signals to test time-varying predictions derived from a computational model. Middle Panel: Multivoxel pattern analysis, in contrast, applies classification techniques to stationary multivariate brain signals (multiple independent instances of a specific pattern of activity across multiple voxels) in order to infer if information about a specific condition or sensory input is encoded in brain activation. Lower panel: Multivariate model-based analysis is a combination of the two: we plan to use computational models of RL to make predictions about the temporal dynamic of information encoded in multivariate brain activity patterns. We then plan to test these predictions using novel multivariate analysis techniques, such as cVMANOVA.

Multivariate fMRI data analysis

Like model-based fMRI analyses, multivariate methods are a relatively recent development in fMRI data analysis that has proven to be very powerful. The commonly used multi-voxel pattern analysis (MVPA) approach, which has emerged from the application of machine learning classification techniques (linear discriminant analysis, support vector machines) to neuroimaging data, has allowed researchers to make inferences about the content of visual input from fMRI data (Haxby et al., 2001), to access information that is represented at spatial scales below fMRI signal's resolution (Haynes and Rees, 2005; Kamitani and Tong, 2005) and to identify abstract mental states (Haynes et al., 2007). In the years following these initial studies, the use of MVPA has vastly increased, leading to a wealth of important discoveries (e.g., Kay et al., 2008). In addition, work in the Haynes lab has shown that even scalar quantities derived from RL theory, such as reward value, have distributed codes on the neural level and thus can be better studied using MVPA (Kahnt et al., 2010, 2011).

The traditional “decoding” approach in MVPA uses the prediction accuracy of a classifier to measure the amount of information that is carried by the fMRI signals (see Figure 1, middle panel). In practice, a classifier is usually trained on a subset of the data (training set) and then tested on the held-out data (test set). This method is designed to avoid spurious results due to over-fitting, however, ***it is not appropriate for testing hypotheses about multivariate representations that change over time, as is necessary in order to test how information is filtered dynamically to allow for efficient RL.*** Indeed, recent evidence by one of the postdocs involved in the project has shown that the basic representation of a task can change through learning, and that these changes take place at different points in time for different individuals (Schuck et al., 2015), thus supporting the need for methods which are better suited for tracking time-varying representations reliably.

The Haynes lab has recently developed a novel MVPA measure, “pattern distinctness” (Allefeld and Haynes, 2014), that is based on a classical multivariate analysis of variance (MANOVA, cf. Timm, 2002) and will provide an excellent starting point for our project. In contrast to classification accuracy, pattern distinctness is a measure of multivariate effect size that works with arbitrary experimental designs including factorial designs and time-varying predictors. By using cross-validation on hold-out (test) sets, Allefeld & Haynes transformed a traditional multivariate test statistic, the Bartlett–Lawley–Hotelling trace, into an unbiased estimator of multivariate effect size that is independent of the partitioning of the data into “training” and “test” parts. The flexibility of cross-validated MANOVA (cvMANOVA) with respect to the experimental designs – allowing one to test hypotheses based on continuous predictors rather than distinct classes of data – is the key feature that allows this method to directly connect RL-based predictions to MVPA. ***cvMANOVA is ideal for combining MVPA and model-based fMRI, because RL-based predictors can be entered into the multivariate analysis similarly to how these predictors are now used in univariate analyses.***

An uncharted territory: Application of multivariate model-based fMRI analysis to reinforcement learning

The aim of the proposed collaboration is to realize the potential of cvMANOVA for advancing our understanding of how learning and decision making occur in complex, multidimensional tasks. In particular, we propose to utilize a novel computational model of RL to make predictions about the temporal dynamics of information encoded in neural representations, and test these predictions using cvMANOVA, while also extending this novel method to specifically address questions about time-varying representations in RL (Figure 1, lower panel). To this end, we will parallelize and intertwine the developments of computational models of RL and novel fMRI analysis methods that have so far been pursued separately in the Niv and Haynes labs. In particular, we propose two specific aims:

1) To apply and test cross-validated MANOVA on existing fMRI datasets from a multidimensional RL task.

In the first aim, we plan to apply the cvMANOVA method, developed in the Haynes lab, to existing fMRI datasets from a multidimensional RL task developed in the Niv lab. With this data set, we can benchmark cvMANOVA against existing methods (classical decoding approaches). Since all components of this aim (i.e., the cvMANOVA method, a computational model of RL that makes predictions about changes in the content of neural representations, and appropriate data from human participants performing the learning task) already exist, the Niv lab can commence this aim immediately with the start of the funding. The results and experiences from this first aim will give us a better understanding of the specific requirements of the planned cvMANOVA extension, which the Haynes lab will develop concurrently (see Aim 2 below).

2) To extend cross-validated MANOVA to non-independent, time-varying data.

cvMANOVA at present only provides a summary measure quantifying the amount of information in a pattern of activity over time. While classification can be applied to single trials, it provides a very crude measure in this case, because the classification of a single trial can only be a hit or a miss. Fortunately, the MANOVA-based approach can be adapted such that it provides a continuous trial-by-trial measure of information content. Realizing this adaptation will be an important aspect of our endeavor. A second aspect will be to extend the scope of cvMANOVA applicability to situations in which representations need to be learned over longer periods of time. This represents a challenge for existing methods because cross validation typically requires that the data be divided into independent, identically distributed subparts. If the cognitive processes in question are developing over time, this will lead to non-identically distributed data. Possible solutions for this problem that will be pursued in this aim are to replace cross-validation by other methods for removing bias in statistical estimators, e.g. the jackknife (Efron and Tibshirani, 1993), permutation-based methods (Good, 2000) or theoretical analysis of the sampling properties of pattern distinctness.

As the Haynes lab develops improved variants of cvMANOVA, the Niv lab will test them on the same datasets, in order to compare each variant against the basic cvMANOVA and against other existing multivariate methods. Ideally, we would also acquire new pilot data that are better suited for testing these methods, as we predict that our existing data will provide a suitably challenging “stress test” for only some components of the cvMANOVA method, and not others (see ideal and minimal budget requirements).

Project Statement

With the current project, the applicants seek to establish a collaboration that will eventually allow them to formulate and test a computational model of reinforcement-based changes in attentional filters and how the computations suggested by the model might be implemented in the human brain. In the service of this basic science goal, we aim to develop a new multivariate model-based fMRI analysis approach that will be of wide use to the field at large.

Contribution to the internationalization of the participating labs

The proposed collaboration will be a seed for a longer-term relationship between the two labs. While our specific aims can be completed with the current grant alone, the applicants intend to develop an extensive research program based on these initial steps. In particular, pending a successful development of the collaboration, the applicants intend to apply for a Collaborative Research in Computational Neuroscience (CRCNS) grant. This multi year grant mechanism is particularly well suited to the proposed work because it is aimed at collaborations that seek to develop computational models and analytical methods that will be tested with experimental approaches. The program is jointly funded by the National Science Foundation and the German Ministry of Education and Research, and thus provides an ideal match for the described US-German collaboration. Total amount for CRCNS grants can be up to 540,000 EUR over a period of 3 years. Applications to this larger grant, however, require substantial preliminary results. We believe that with the help of the Humboldt-Princeton funding, a strong application for a CRCNS grant can be prepared.

Involvement of Princeton and Humboldt-Universität scholars

The above-mentioned work on computational models of RL in Yael Niv’s lab and the work on extensions of multivariate analysis methods in John-Dylan Haynes’s lab will serve as the foundation of this collaboration. The expertise of the two groups is highly complementary, which promises a fruitful synergy. This expertise and the fact that initial data and methods exist will allow us to begin the project from day one (see timetable below).

The proposed collaboration is entirely new and has the potential to lead to novel, publishable findings, methods that can be applied broadly, and preliminary data that

will support a CRCNS grant proposal to extend the collaboration beyond the initial Humboldt-Princeton grant. Even though the two labs have not collaborated before, the lead applicants know each other personally, value each others work and therefore consider it a serendipity that their host institutes have set up a grant specifically tailored for Humboldt-Princeton collaborations.

In addition to the lead applicants, the following people will be involved in the collaboration:

- **Dr. Nicolas Schuck**, senior postdoctoral researcher in the Niv Lab, has been involved in the development of computational models of RL in the Niv lab over the past two years. Dr. Schuck also has considerable expertise in MVPA methods. As part of the proposed collaboration, Dr. Schuck will visit the Haynes Lab for an extended period, where he will get further training in cvMANOVA methods. Importantly, Dr. Schuck will provide a model-based fMRI perspective to the ongoing development of new analysis methods in the Haynes Group.
- **Dr. Carsten Allefeld**, postdoctoral researcher in the Haynes Lab, has been working on the development of multivariate analysis methods for fMRI. Dr. Allefeld will be the main collaboration partner of Dr. Schuck and will work on the adaptation of cvMANOVA methods to fit the specific needs of model-based fMRI of RL processes. As part of the proposed collaboration, Dr. Allefeld will visit the Niv Lab for an extended period, where he will get further training in RL methods.
- **Ms. Angela Radulescu**, PhD student in the Niv lab, is involved in the development of models of RL in multidimensional environments. Within the proposed collaboration, Ms. Radulescu will be involved in the ongoing efforts to develop a multivariate model-based fMRI approach and will receive further training in matters of computational modeling as well as univariate and multivariate fMRI data analysis through her work with Dr. Schuck and through a visit to the Haynes lab.